Field theoretic approach to the counting problem of Hamiltonian cycles of graphs

Saburo Higuchi*
Department of Pure and Applied Sciences,
The University of Tokyo
Komaba, Meguro, Tokyo 153-8902, Japan

(November 14, 1997, cond-mat/9711152, UT-KOMABA/97-12, Revised on February 3, 1998)

A field theoretic representation for the number of Hamiltonian cycles of graphs is studied. By integrating out quadratic fluctuations around the saddle point, one obtains an estimate for the number which reflects characteristics of graphs well. The accuracy of the estimate is verified by applying it to 2d square lattices with various boundary conditions. This is the first example of extracting meaningful information from the quadratic approximation to the field theory representation.

I. INTRODUCTION

Let $G = (V, E)$ be a graph with the set of vertices $V = \{r_j\}$ and of the edges $E = \{e_k\}$. A Hamiltonian cycle of a graph is a closed path which visits each of the vertices once and only once. I denote the number of all the Hamiltonian cycles of a graph G by $H(G)$:

$$H(G) = \sum_{\text{Hamiltonian cycle on } G} 1. \tag{1}$$

See Fig.1 for examples.

Hamiltonian cycles have often been used to model collapsed polymer globules [1]. The quantity $H(G)$ corresponds to the entropy of a polymer system in G in a collapsed but disordered phase. One can model even more realistic polymers by introducing a weight that depends on the shape of cycles in (1). The polymer melting problem is studied by taking into account the bending energy with a weight which depends on the number of turns on the cycle [2]. In ref. [3], the protein folding problem is studied by incorporating the Van der Waals potential as well as the bending energy in the model (1).

For homogeneous graphs (lattices) with N vertices, one expects that $H(G)$ behaves like

$$H(G) \to C(G) N^{\gamma-1} \omega^N \qquad (N \to \infty), \tag{2}$$

where ω is defined by

$$\log \omega = \lim_{N\to\infty} \frac{1}{N} \log H(G). \tag{3}$$

The quantity ω is supposed to be a universal bulk quantity whereas $C(G)$ and γ depend on the detail of graphs *e.g.* boundary conditions [4].

A field theory representation for (1) for arbitrary graphs is introduced in ref. [5] and has been used to study the extended models [2,3]. For homogeneous graphs with the number of vertices N and the coordination number q, the saddle point approximation to the representation yields

$$H(G) \simeq \left(\frac{q}{e}\right)^N \tag{4}$$

or $\omega \simeq q/e$. This approximation has been proved to be good in many examples [6]. For a square lattice, $q/e = 4/e = 1.4715\ldots$ is quite near to the exact value $\omega \simeq 1.473$ estimated by the direct enumeration [7] and other methods [4,8,9].

In the saddle point approximation, however, graphs with identical (q, N) are not distinguished. Indeed there is a variety of graphs which has (q, N) in common. Given a graph with a pair (q, N), it is often possible to change its boundary conditions to modify the 'topology' and the 'moduli' of the graph keeping the pair. Fig. 1 shows examples. Moreover, there are graphs which have identical (q, N) but have distinct local structures, *e.g.* the 2d triangular lattice and the 3d cubic lattice.

In this article, I go beyond the saddle point approximation. I work out the quadratic approximation and find an estimate for $H(G)$ whose G-dependence is not merely through (q, N) but is more sensible. To demonstrate the validity of the approximation, I examine 2d square lattices with a variety of boundary conditions and aspect ratios. It is also examined if the estimate $\omega \simeq q/e$ is improved.

II. FIELD THEORETIC REPRESENTATION

The problem of calculating $H(G)$ is mapped into one in a lattice field theory on G [5]. This is done by introducing an O(n) lattice field $\vec{\phi}(r) = (\phi_1(r), \ldots, \phi_n(r))$ living on $V = \{r_j\}$ with an action

$$S[\vec{\phi}(r)] = \frac{1}{2} \sum_{r,r' \in V, 1\le j \le n} \phi_j(r)(i\Delta^{-1} + \epsilon)_{rr'} \phi_j(r'). \tag{5}$$

The $N \times N$ matrix Δ is the adjacency matrix [10] of the graph G:

$$\Delta_{rr'} = \begin{cases} 1 & \text{if } r, r' \in V \text{ is connected by an } e \in E \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

and an infinitesimal parameter $\epsilon > 0$ is introduced for convergence.

The integer $H(G)$ is related to a $2N$-point function by

$$H(G) = \lim_{n\to 0}\lim_{\epsilon\to +0}\frac{1}{n}\left|\frac{Z_1}{Z_0}\right|, \tag{7}$$

$$Z_0 = \int D\vec{\phi} \quad \mathrm{e}^{-S[\vec{\phi}(r)]}, \tag{8}$$

$$Z_1 = \int D\vec{\phi} \quad \mathrm{e}^{-S[\vec{\phi}(r)]} \prod_{r\in V}\frac{\vec{\phi}^2(r)}{2}, \tag{9}$$

where $D\vec{\phi} = \prod_{r\in V, 1\le j\le n} d\phi_j(r)$. Eq. (7) holds for arbitrary G with $N > 2$ because each terms in the diagrammatic expansion corresponds to a Hamiltonian cycle.

III. APPROXIMATION

First I evaluate (7) by the saddle point method. I concentrate on Z_1 since $Z_0 \to 1$ as $n \to 0$. When a graph G is homogeneous, there are mean field saddle points which are degenerate on

$$\{\vec{\phi}|\ \vec{\phi}^2(r) \equiv -2iq_\epsilon\} \simeq \frac{\mathrm{O}(n)}{\mathrm{O}(n-1)}, \tag{10}$$

where

$$\frac{1}{q_\epsilon} := \frac{1}{q} - i\epsilon. \tag{11}$$

This yields the estimate (4).

Then I consider fluctuations around the saddle point. It can easily be seen that there are zero modes corresponding to the global $\mathrm{O}(n)/\mathrm{O}(n-1)$ symmetry and sublattice symmetries [5]. Therefore I am led to introduce a gauge fix condition

$$\sum_{r\in V}\phi_j(r) = 0 \qquad (2 \le j \le n). \tag{12}$$

By the standard Fadeev-Popov method, I have

$$Z_1 = \int D\phi\, e^{-\frac{1}{2}\sum_{r,r'\in V, 1\le j\le n}\phi_j(r)(i\Delta^{-1}+\epsilon)_{rr'}\phi_j(r')}\left|\sum_{r\in V}\phi_1(r)\right|^{n-1}\frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}\prod_{2\le j\le n}\delta\left(\sum_{r\in V}\phi_j(r)\right)\prod_{r\in V}\frac{\vec{\phi}^2(r)}{2}. \tag{13}$$

The factor $|\sum_r \phi_1(r)|^{n-1}\pi^{\frac{n}{2}}/\Gamma(\frac{n}{2})$ is the Fadeev-Popov determinant multiplied by the volume of the gauge orbit. Actually, this is nothing but the jacobian for the n-dimensional radial coordinate for constant modes.

Expanding the field $\vec{\phi}$ around the saddle point as

$$\phi_j(r) = \sqrt{-2iq_\epsilon I_{N,n}}\,\delta_{j1} + \psi_j(r), \tag{14}$$

where

$$I_{N,n} = 1 + \frac{n-1}{2N}, \tag{15}$$

one obtains

$$Z_1 = \left(\frac{q_\epsilon I_{N,n}}{ie}\right)^{NI_{N,n}} 2^{\frac{n-1}{2}} N^{n-1}\frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}$$
$$\times \int D\vec{\psi}\, e^{-\frac{1}{2}\sum_{r,r'\in V}\psi_1(r)A^{\mathrm{L}}_{rr'}(n)\psi_1(r')} \times e^{-\frac{1}{2}\sum_{r,r'\in V, 2\le j\le n}\psi_j(r)A^{\mathrm{T}}_{rr'}(n)\psi_j(r')} \times e^{-V_{\mathrm{int}}(\vec{\psi})} \times \prod_{2\le j\le n}\delta\left(\sum_{r\in V}\psi_j(r)\right), \tag{16}$$

where $A^{\mathrm{L}}(n)$ and $A^{\mathrm{T}}(n)$ are the inverse propagators of the longitudinal mode ($j = 1$) and the transverse modes ($2 \le j \le n$), respectively:

$$A^{\mathrm{L}}_{rr'}(n) = i(\Delta^{-1})_{rr'} + \frac{i}{q}I^{-1}_{N,n}\delta_{rr'} + (1 + I^{-1}_{N,n})\epsilon\delta_{rr'} + \frac{1}{N}\frac{i}{q_\epsilon}(1 - I^{-1}_{N,n}), \tag{17}$$

$$A^{\mathrm{T}}_{rr'}(n) = i(\Delta^{-1})_{rr'} - \frac{i}{q}I^{-1}_{N,n}\delta_{rr'} + (1 - I^{-1}_{N,n})\epsilon\delta_{rr'}. \tag{18}$$

One neglects the interaction terms $V_{\mathrm{int}}(\vec{\psi})$ and performs the gaussian integrations to obtain

$$Z_1 \simeq \left(\frac{qI_{N,n}}{ie}\right)^{NI_{N,n}} (2N)^{\frac{n-1}{2}}\frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})}\left[\frac{(2\pi)^{\frac{N-1}{2}}}{\det'^{\frac{1}{2}} A^{\mathrm{T}}(n)}\right]^{n-1}\frac{(2\pi)^{\frac{N}{2}}}{\det^{\frac{1}{2}} A^{\mathrm{L}}(n)} \times 2, \tag{19}$$

where ϵ is sent to $+0$. The last factor two corresponds to the residual symmetry $\sum_r \phi_1(r) \leftrightarrow -\sum_r \phi_1(r)$. The prime on det means the omission of the eigenvalue for the constant mode.

One notices that the signature of the real part (the last term) of (18) changes at $n=1$. Thus one has to assume $n>1$ to derive (19). Then in eq.(19) one takes the limit $n \to 0$ to obtain the final result

$$H(G) \simeq \left(\frac{q}{e}\right)^{N} e^{\frac{1}{2}} (I_{N,0})^{N I_{N,0}} \sqrt{\frac{\pi}{N}} \frac{\det'^{\frac{1}{2}}(A^{\mathrm{T}}(0)\Delta)}{\det^{\frac{1}{2}}(A^{\mathrm{L}}(0)\Delta)}. \quad (20)$$

Note that $\det\Delta \det' A^{\mathrm{T}}(0) = q \det'(A^{\mathrm{T}}(0)\Delta)$. Eq. (20) is nothing but the estimate I desired to have. The saddle point result $(q/e)^N$ is corrected by the ratio of determinants which contains information of details of the structure of G.

In ref. [5], it is claimed that the quadratic correction to (4) vanishes. In the present analysis, the Fadeev-Popov method is worked out to find $\sqrt{\pi/N}$ missing in ref. [5]. As shown in Sec.IV, the ratio of determinants is not equal to unity and contributes to γ and $C(G)$ non-trivially. Moreover, inclusion of $i=\sqrt{-1}$ and $\epsilon>0$ in the action (5) in the present analysis enables one to discuss the limit of application of the quadratic approximation in Sec.IV.

IV. SQUARE LATTICES

To see how eq.(20) works, I study concrete examples $P(L_1,L_2)$ and $SP(L_1,L_2)$ shown in Fig.1. Both are two-dimensional square lattices with the edge lengths L_1 and L_2. The difference between P and SP lies in boundary conditions. For $P(L_1,L_2)$, the periodic boundary condition is imposed for both two directions. For $SP(L_1,L_2)$, that across the edge L_2 is replace by the skew-periodic one. They are good examples to test eq.(20). One can switch the boundary condition or vary the aspect ratio L_2/L_1 to make the graph globally distinct while keeping $(q,N)=(4,L_1\times L_2)$. The saddle point approximation cannot see the difference among them but eq.(20) has a chance to distinguish them.

Graphs $P(L_1,L_2)$ and $SP(L_1,L_2)$ can be viewed as discrete tori with different moduli parameters. Thus it is interesting also in its own right to determine the asymptotic behaviors of $H(P(L_1,L_2))$ and $H(SP(L_1,L_2))$ in the limit $L_1,L_2\to\infty$.

I analytically evaluate (20) for $P(L_1,L_2)$ and $SP(L_1,L_2)$. In the momentum representation, the determinants become ($\sigma=0$ for P and $\sigma=1$ for SP)

$$\det{}'(A^{\mathrm{T}}(n)\Delta) = \quad {\prod}'_{0\le n_j\le L_j-1}\left[1-\frac{1}{2I_{N,n}}\left(\cos\left(k_1+\sigma\frac{k_2}{L_1}\right)+\cos k_2\right)\right], \quad (21)$$

$$\det(A^{\mathrm{L}}(n)\Delta) = 2\times{\prod}'_{0\le n_j\le L_j-1}\left[1+\frac{1}{2I_{N,n}}\left(\cos\left(k_1+\sigma\frac{k_2}{L_1}\right)+\cos k_2\right)\right]. \quad (22)$$

The constant mode $k_1=k_2=0$ is excluded in $\prod'$. Hereafter, the indices n_j and k_j should be related by $k_j=2\pi n_j/L_j$.

The formula

$$\prod_{0\le m\le L-1}\left[x^2-2x\cos\left(\theta+\frac{2m\pi}{L}\right)+1\right]=x^{2L}-2x^L\cos(L\theta)+1 \quad (23)$$

enables one to obtain

$$\frac{\det'(A^{\mathrm{T}}(0)\Delta)}{\det(A^{\mathrm{L}}(0)\Delta)} = -2L_1L_2 I_{L_1L_2,0}\prod_{n_2=0}^{L_2-1}\frac{u_-(k_2)}{u_+(k_2+\pi)}, \quad (24)$$

where

$$u_\pm(k)=\begin{cases} y(k)^{L_1}+y(k)^{-L_1}\pm 2, & \text{for } P,\\ y(k)^{L_1}+y(k)^{-L_1}-2\cos k & \text{for } SP,\end{cases} \quad (25)$$

$$y(k)=2I_{L_1L_2,0}-\cos k+\sqrt{(2I_{L_1L_2,0}-\cos k)^2-1}. \quad (26)$$

Now I specialize to the case where both L_1 and L_2 are odd. In the limit $N=L_1\times L_2\to\infty$ with $R=L_2/L_1$ fixed, it is allowed to approximate $\prod_{n_2}$ by $\exp(\int dk_2 \log)$ after separating quickly oscillating factors. In this limit, assuming $0<R\le 1$ for periodic case, I obtain

$$H(P(L_1,L_2)) \simeq \left(\frac{q}{e}\right)^N \sqrt{\frac{\pi}{2}}\frac{\sin(\frac{1}{2R})^{\frac{1}{2}}\exp(\frac{9\pi^2}{16R^2}-\frac{1}{2R})^{\frac{1}{2}}}{\cosh^2(\frac{\pi^2}{4R^2}-\frac{1}{2R})^{\frac{1}{2}}} \quad (27)$$

$$H(SP(L_1,L_2)) \simeq \left(\frac{q}{e}\right)^N \sqrt{2\pi}\frac{\sin(\frac{1}{2R})^{\frac{1}{2}}\left|\sinh(\frac{9\pi^2}{16R^2}-\frac{1}{2R})^{\frac{1}{2}}\right|}{\left|\sinh^2(\frac{\pi^2}{4R^2}-\frac{1}{2R})^{\frac{1}{2}}\right|} \quad (28)$$

One sees the quadratic approximation predicts the explicit form of $C(G)$ and the fact $\gamma=1$ in (2). It is remarkable that the correction is independent of N in the limit. Namely, it implies that the estimates for ω are unchanged.

Figs 2 and 3 show the plot of (27), (28) normalized by $(4/e)^{L_1L_2}$ as functions of L_2/L_1 (solid line). The finite L_1L_2 results (20) for odd L_j such that $3\le L_j\le 29$ are also shown (solid circle). For the purpose of comparison, I plot the exact numbers of Hamiltonian cycles (box) and estimates of them by simulations (solid box) for L_j odd [11].

The exact numbers are determined by direct enumeration by a program in C. Table I shows the result.

This program is useful only for small graphs because the time needed for calculation increases exponentially with $L_1 \times L_2$. The simulation is based on the biased Monte Carlo method with a code in C. The time needed grows again exponentially with $L_1 \times L_2$ but with a smaller exponent. I have used direct enumeration for $L_1 \times L_2 \leq 39$ and Monte Carlo simulation for $L_1 \times L_2 \lesssim 90$. The evaluation of (20) takes time proportional to $L_1 \times L_2$. Therefore eq.(20) has an advantage even if it is approximate.

Figs 2 and 3 suggest that the quadratic approximation to the field theory is reliable. The field theory succeeds in predicting that the correction depends almost only on L_2/L_1 and that there is a qualitative difference between P and SP. The L_2/L_1-dependence of the correction generally agrees with the exact result though there is a slight deviation in the small L_2/L_1 region.

There is a definite discrepancy at $L_2/L_1 = \pi^2/2$ for the skew-periodic case. The quadratic approximation (28) diverges at $L_2/L_1 = \pi^2/2$ as seen in Fig.3 while the exact result takes finite values and is simply increasing there. Actually, one can argue that $L_2/L_1 \gtrsim \pi^2/2$ is out of range of application of (20). As I mentioned, I defined the limit $n \to 0$ as the continuation from $n > 1$ where the gaussian integral converges. Let us look at the evolution of the spectrum of $(A^{\mathrm{L}}(n)\Delta)$ on the way from $n > 0$ down to $n = 0$. The eigenvalue for $k_j = (1-1/L_j)\pi$ mode hits zero at some $0 < n_{\mathrm{c}} < 1$ if

$$2I_{L_1 L_2,0} < \cos\frac{\pi}{L_1 L_2} + \cos\frac{\pi}{L_2}. \tag{29}$$

This condition is equivalent to $L_2/L_1 > \pi^2/2$ in the limit $L_1 \times L_2 \to \infty$. So one suddenly has a zero mode at $n = n_{\mathrm{c}}$ and no symmetry is responsible for it. This suggests that the quadratic approximation breaks down there and that (28) is not reliable for such values of L_2/L_1. A careful analysis shows that the periodic case is free from such spurious zero modes.

V. SUMMARY AND DISCUSSIONS

I have found an approximate formula for the number of Hamiltonian cycles of graphs by the quadratic approximation to the field theoretic representation. It has been tested for 2d square lattice with a variety of boundary conditions. I have obtained a multiplicative correction which is independent of the size $N = \#V$. Namely, the dependence on the boundary conditions is obtained and is shown to be in a good agreement with the true behavior.

A natural extension to the present problem is the counting of the number of closed self-avoiding walks that visit fN vertices of a graph, where $0 \leq f \leq 1$ is fixed. In that case, the form (2) can still be assumed but $C(G), \gamma$ and ω now depend on f. There is a field theoretic representation for $\omega(f)$ called lattice cluster theory [12]. It is not identical with (7) even for $f = 1$. It is interesting to compare the quadratic approximation to lattice cluster theory [13,14] with the present result. In the former one, estimate for $\omega(f)$ is improved over the saddle point approximation for $f < 1$. For $f = 1$, the estimate for $\omega = \omega(1)$ is unchanged in accord with the present case.

ACKNOWLEDGMENTS

I thank T. Yamaguchi for collaboration in the initial stage of this work. Many helpful conversations with S. Hikami, M. Itakura, K. Minakuchi, R. Pnini and J. Suzuki are gratefully acknowledged. This work is partially supported by Grant-in-Aid's for Scientific Research (No. 08740196 and No. 08454106) from the Ministry of Education, Science and Culture, and by CREST from Japan Science and Technology Corporation.

* Electronic address: hig@rice.c.u-tokyo.ac.jp

[1] J. des Cloizeaux and G. Jannik, *Polymers in solution: their modelling and structure* (Clarendon Press, Oxford, 1987).

[2] J. Bascle, T. Garel, and H. Orland, J. Phys. A Math. Gen. **25**, L1323 (1992).

[3] S. Doniach, T. Garel, and H. Orland, J. Chem. Phys. **105**, 1601 (1996).

[4] T. G. Schmalz, G. E. Hite, and D. J. Klein, J. Phys. A Math. Gen. **17**, 445 (1984).

[5] H. Orland, C. Itzykson, and C. de Dominicis, J. Physique **46**, L353 (1985).

[6] J. Suzuki, J. Phys. Soc. Japan **57**, 687 (1988).

[7] C. J. Camacho and D. Thirumalai, Phys. Rev. Lett. **71**, 2505 (1993).

[8] M. T. Batchelor, H. W. J. Blöte, B. Nienhuis, and C. M. Yung, J. Phys. A Math. Gen. **29**, L399 (1996).

[9] B. Duplantier and H. Saleur, Nucl. Phys. **B290[FS20]**, 291 (1987).

[10] In the case $\det\Delta = 0$, the action (5) is ill-defined. In that case I can justify (7) by replacing Δ with, for example,

$$\Delta(\epsilon_0)_{rr'} = \Delta_{rr'} + \epsilon_0\delta_{rr'}, \tag{30}$$

where ϵ_0 is a small number. Then I have $\det(\Delta(\epsilon_0)) \neq 0$ and the field theory yields

$$\langle\phi_j(r)\phi_k(r')\rangle = \delta_{jk}\Delta(\epsilon_0)_{rr'} \to \delta_{jk}\Delta_{rr'} \quad (\epsilon_0 \to 0) \tag{31}$$

as desired. So eq.(7) is valid with $\epsilon_0 \to 0$ limit understood. By this regularization, the estimates (4) and (20), which are free from $\Delta^{-1} = \Delta(0)^{-1}$, remain valid.

[11] For the case with L_j even, the L_2/L_1-dependence does not seem to be the same. Work on this point is in progress.

[12] K. F. Freed, J. Phys. A Math. Gen. **18**, 871 (1985).

[13] A. M. Nemirovsky and M. D. Coutinho-Filho, J. Statist. Phys. **53**, 1139 (1988).

[14] A. M. Nemirovsky and M. D. Coutinho-Filho, Phys. Rev. **A39**, 3120 (1989).

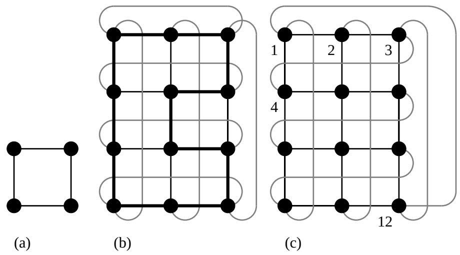

FIG. 1. Examples of graphs and Hamiltonian cycles on them. One does not distinguish the base points and the directions of a cycle. (a) $H(G) = 1$ for this graph. (b) $P(3, 4)$, the 2d square lattice with the periodic boundary condition. A Hamiltonian cycle is drawn in the thick line. For this graph, $(q, N) = (4, 12)$. (c) $SP(3, 4)$, the 2d square lattice with the skew-periodic boundary in the horizontal direction. It is locally isomorphic to $P(3, 4)$ but the boundary condition makes it distinct globally.

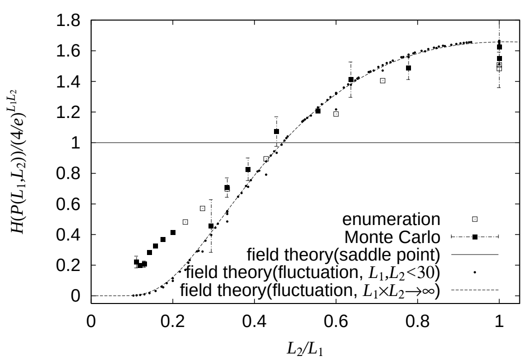

FIG. 2. The number of Hamiltonian cycles for $P(L_1, L_2)$ for odd L_j as a function of L_2/L_1. Normalized by the saddle point result $(4/e)^{L_1 L_2}$. Plotted are the quadratic approximation to the field theory (solid square: $3 \leq L_1, L_2 \leq 29$, dashed line: the limit $L_1 \times L_2 \to \infty$), exact results obtained by enumeration (box, $L_1 \times L_2 \leq 39$), and estimates by weighted Monte Carlo simulations (solid box, $L_1 \times L_2 \lesssim 90$).

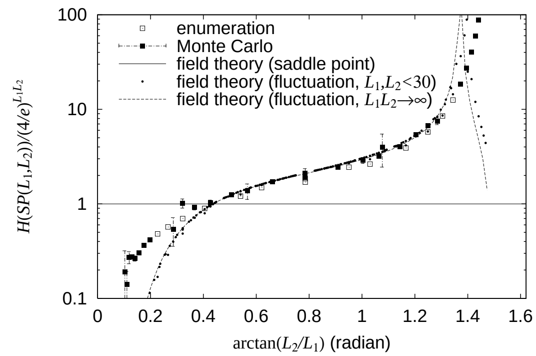

FIG. 3. The log plot of the number of Hamiltonian cycles for $SP(L_1, L_2)$ for odd L_j as a function of $\arctan(L_2/L_1)$. Normalized by the saddle point result $(4/e)^{L_1 L_2}$. Plotted are the same as in Fig.3.

TABLE I. Exact number of Hamiltonian cycles for $L_1 \times L_2$ 2d square lattice with periodic (P) and skew-periodic (SP) boundary conditions. Determined by the direct enumeration.

L_1	L_2	$H(P(L_1, L_2))$	$H(SP(L_1, L_2))$	$H(SP(L_2, L_1))$
3	3	48	55	55
5	3	390	397	866
5	5	23580	29001	29001
7	3	2982	2989	13021
7	5	1045940	1108006	1820582
9	3	23646	23653	195157
11	3	196086	196093	2924373
13	3	1682382	1682389	43820323